

RAW SEQUENCE LISTING  
PATENT APPLICATION US/09/986,682A

2/04

DATE: 02/21/2002  
TIME: 22:54:14

INPUT SET: S36770.raw

#4

This Raw Listing contains the General  
Information Section and up to the first 5 pages.

SEQUENCE LISTING

- 1
- 2
- 3 (1) General Information:
- 4 (i) APPLICANT: Tsusaki et al.
- 5 (ii) TITLE OF INVENTION: POLYPEPTIDE HAVING -FRUCTOFURANOSIDASE
- 6 ACTIVITY
- 7 (iii) NUMBER OF SEQUENCES: 6
- 8 (iv) CORRESPONDENCE ADDRESS:
- 9 (A) ADDRESSEE: BROWDY AND NEIMARK
- 10 (B) STREET: 419 Seventh Street, N.W., Suite 300
- 11 (C) CITY: Washington
- 12 (D) STATE: D.C.
- 13 (E) COUNTRY: USA
- 14 (F) ZIP: 20004
- 15 (v) COMPUTER READABLE FORM:
- 16 (A) MEDIUM TYPE: Floppy disk
- 17 (B) COMPUTER: IBM PC compatible
- 18 (C) OPERATING SYSTEM: PC-DOS/MS-DOS
- 19 (D) SOFTWARE: PatentIn Release #1.0, Version #1.30
- 20 (vi) CURRENT APPLICATION DATA:
- 21 (A) APPLICATION NUMBER: 09/986,682
- 22 (B) FILING DATE:
- 23 (C) CLASSIFICATION:
- 24 (vii) PRIOR APPLICATION DATA:
- 25 (A) APPLICATION NUMBER: US/09/317,179
- 26 (B) FILING DATE:
- 27 (viii) ATTORNEY/AGENT INFORMATION:
- 28 (A) NAME: BROWDY, Roger L.
- 29 (B) REGISTRATION NUMBER: 25,618
- 30 (C) REFERENCE/DOCKET NUMBER: TSUSAKI=2
- 31 (ix) TELECOMMUNICATION INFORMATION:
- 32 (A) TELEPHONE: 202-628-5197
- 33 (B) TELEFAX: 202-737-3528
- 34 (2) INFORMATION FOR SEQ ID NO: 1:
- 35 (i) SEQUENCE CHARACTERISTICS:
- 36 (A) LENGTH: 21 amino acids
- 37 (B) TYPE: amino acid
- 38 (D) TOPOLOGY: linear
- 39 (v) FRAGMENT TYPE: N-terminal fragment
- 40 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1:
- 41 Met Asn Ser Gly Asp Tyr Lys Glu Asp Tyr Gly Phe Ala His Ile Thr
- 42 1 5 10 15
- 43 Arg Ala Asp Met Leu
- 44 20
- 45
- 46 (2) INFORMATION FOR SEQ ID NO: 2:

ENTERED

# RAW SEQUENCE LISTING PATENT APPLICATION US/09/986,682A

DATE: 02/21/2002  
TIME: 22:54:15

INPUT SET: S36770.raw

```

47      (i) SEQUENCE CHARACTERISTICS:
48          (A) LENGTH: 12 amino acids
49          (B) TYPE: amino acid
50          (D) TOPOLOGY: linear
51      (v) FRAGMENT TYPE: internal fragment
52      (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2:
53      Ser Val Phe Asp Gly Gly Asp Gly Thr Val Tyr Gln
54      1          5          10
55
56      (2) INFORMATION FOR SEQ ID NO: 3:
57          (i) SEQUENCE CHARACTERISTICS:
58              (A) LENGTH: 455 amino acids
59              (B) TYPE: amino acid
60              (D) TOPOLOGY: linear
61          (ii) MOLECULE TYPE: peptide
62          (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 3:
63      Met Asn Ser Gly Asp Tyr Lys Glu Asp Tyr Gly Phe Ala His Ile Thr
64      1          5          10          15
65      Arg Ala Asp Met Leu Lys Ile Pro Gly Gln Gln Asn Ser Pro Gln Phe
66      20          25          30
67      Lys Val Pro Gln Phe Asn Ala Ser Ala Ile Lys Asn Ile Asp Ser Ala
68      35          40          45
69      Lys Gly Tyr Asp Lys Ser Gly Asn Leu Ile Asp Leu Asp Val Trp Asp
70      50          55          60
71      Ser Trp Pro Leu Gln Asn Ala Asp Gly Thr Ala Ala Asn Tyr His Gly
72      65          70          75          80
73      Tyr His Ile Val Ser Ala Leu Ala Gly Asp Pro Lys Asn Ser Asp Asp
74      85          90          95
75      Thr Pro Leu His Leu Phe Tyr Gln Lys Val Gly Asp Thr Ser Ile Asp
76      100          105          110
77      Ser Trp Lys Asn Ala Gly Arg Val Phe Glu Asp Met Asp Lys Phe Val
78      115          120          125
79      Pro Asn Asp Pro Tyr Leu Lys Tyr Gln Thr Gln Glu Trp Ser Gly Ser
80      130          135          140
81      Ala Thr Leu Thr Lys Asp Gly Gln Val Arg Leu Phe Tyr Thr Asp Tyr
82      145          150          155          160
83      Ser Gly Asn Pro Glu Asp Gly Gly Thr Gly Ala Gly Asn Gln Ile Ile
84      165          170          175
85      Ser Thr Ala Gln Val Asn Leu Ser Gln Pro Asp Ala Ala Thr Leu Lys
86      180          185          190
87      Val Asp Gly Val Ser Asp His Lys Ser Val Phe Asp Gly Gly Asp Gly
88      195          200          205
89      Thr Val Tyr Gln Asn Ile Gln Gln Phe Ile Asp Glu Gly Lys Trp Ile
90      210          215          220
91      Ser Gly Asp Asn His Thr Leu Arg Asp Pro His Tyr Val Glu Asp Lys
92      225          230          235          240
93      Gly His Lys Tyr Leu Val Phe Glu Ala Asn Thr Gly Thr Thr Asp Gly
94      245          250          255
95      Tyr Gln Gly Asp Gln Ser Phe Asn Asn Lys Ala Tyr Tyr Gly Gly Ser
96      260          265          270
97      Asp Val Phe Phe Gln Asn Glu Lys Asn Lys Leu Leu Gln Ser Pro Lys
98      275          280          285
99      Lys Gln Ile Ala Ser Leu Ala Asn Gly Ala Leu Gly Ile Val Glu Leu

```

# RAW SEQUENCE LISTING PATENT APPLICATION US/09/986,682A

DATE: 02/21/2002  
TIME: 22:54:15

INPUT SET: S36770.raw

100		290		295		300	
101	Ala	Asp	Asp	Tyr	Thr	Val	Lys
102	305					310	
103	Asn	Thr	Val	Ala	Asp	Glu	Val
104						325	
105	Asn	Lys	Trp	Tyr	Leu	Phe	Thr
106						340	
107	Asp	Gly	Ile	Asn	Asp	Lys	Asp
108						355	
109	Ser	Leu	Asn	Gly	Pro	His	Asn
110						370	
111	Asn	Met	Asn	Leu	Asp	Pro	Ala
112	385					390	
113	Gly	Ile	Pro	His	Pro	Glu	Gly
114						405	
115	Thr	Asn	Arg	Gly	Phe	Tyr	Pro
116						420	
117	Leu	Gly	Val	Asn	Ile	Lys	Gly
118						435	
119	Ser	Gly	Gln	Gly	Gln	Phe	Pro
120						450	

(2) INFORMATION FOR SEQ ID NO: 4:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 1365 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: double

(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 4:

129	ATG	AAC	AGC	GGG	GAC	TAC	AAG	GAA	GAC	TAT	GGT	TTT	GCC	CAT	ATT	ACA	48
130	Met	Asn	Ser	Gly	Asp	Tyr	Lys	Glu	Asp	Tyr	Gly	Phe	Ala	His	Ile	Thr	
131	1			5						10					15		
132	CGC	GCT	GAC	ATG	CTA	AAA	ATT	CCA	GGA	CAA	CAA	AAC	AGT	CCT	CAA	TTT	96
133	Arg	Ala	Asp	Met	Leu	Lys	Ile	Pro	Gly	Gln	Gln	Asn	Ser	Pro	Gln	Phe	
134				20						25					30		
135	AAA	GTG	CCT	CAA	TTC	AAT	GCA	TCA	GCA	ATC	AAA	AAC	ATT	GAT	TCG	GCA	144
136	Lys	Val	Pro	Gln	Phe	Asn	Ala	Ser	Ala	Ile	Lys	Asn	Ile	Asp	Ser	Ala	
137				35						40					45		
138	AAA	GGG	TAT	GAT	AAG	TCA	GGC	AAC	TTA	ATA	GAT	TTA	GAT	GTA	TGG	GAT	192
139	Lys	Gly	Tyr	Asp	Lys	Ser	Gly	Asn	Leu	Ile	Asp	Leu	Asp	Val	Trp	Asp	
140				50						55					60		
141	AGC	TGG	CCA	CTG	CAA	AAC	GCT	GAT	GGT	ACT	GCG	GCA	AAT	TAT	CAT	GGA	240
142	Ser	Trp	Pro	Leu	Gln	Asn	Ala	Asp	Gly	Thr	Ala	Ala	Asn	Tyr	His	Gly	
143				65						70					75		
144	TAT	CAC	ATC	GTC	TCC	GCT	TTA	GCA	GGT	GAC	CCA	AAA	AAC	AGT	GAT	GAT	288
145	Tyr	His	Ile	Val	Ser	Ala	Leu	Ala	Gly	Asp	Pro	Lys	Asn	Ser	Asp	Asp	
146				85						90					95		
147	ACT	CCA	CTT	CAT	TTA	TTC	TAT	CAA	AAA	GTC	GGT	GAT	ACA	TCG	ATT	GAC	336
148	Thr	Pro	Leu	His	Leu	Phe	Tyr	Gln	Lys	Val	Gly	Asp	Thr	Ser	Ile	Asp	
149				100						105					110		
150	AGC	TGG	AAA	AAT	GCT	GGA	AGA	GTA	TTT	GAA	GAT	ATG	GAT	AAA	TTT	GTT	384
151	Ser	Trp	Lys	Asn	Ala	Gly	Arg	Val	Phe	Glu	Asp	Met	Asp	Lys	Phe	Val	
152						120				125					130		

# RAW SEQUENCE LISTING PATENT APPLICATION US/09/986,682A

DATE: 02/21/2002  
TIME: 22:54:16

## INPUT SET: S36770.raw

153	CCA	AAT	GAT	CCG	TAT	CTT	AAA	TAT	CAA	ACA	CAG	GAG	TGG	TCA	GGT	TCT	432
154	Pro	Asn	Asp	Pro	Tyr	Leu	Lys	Tyr	Gln	Thr	Gln	Glu	Trp	Ser	Gly	Ser	
155				135					140					145			
156	GCT	ACT	TTA	ACC	AAA	GAT	GGC	CAA	GTC	CGT	TTA	TTC	TAT	ACA	GAT	TAC	480
157	Ala	Thr	Leu	Thr	Lys	Asp	Gly	Gln	Val	Arg	Leu	Phe	Tyr	Thr	Asp	Tyr	
158				150				155					160				
159	TCA	GGT	AAT	CCT	GAA	GAT	GGT	GGA	ACC	GGT	GCT	GGT	AAC	CAA	ATC	ATT	528
160	Ser	Gly	Asn	Pro	Glu	Asp	Gly	Gly	Thr	Gly	Ala	Gly	Asn	Gln	Ile	Ile	
161				165				170					175				
162	TCA	ACT	GCT	CAA	GTA	AAC	TTA	TCC	CAG	CCG	GAT	GCA	GCT	ACA	CTT	AAA	576
163	Ser	Thr	Ala	Gln	Val	Asn	Leu	Ser	Gln	Pro	Asp	Ala	Ala	Thr	Leu	Lys	
164				180				185				190				195	
165	GTC	GAT	GGA	GTA	TCT	GAT	CAT	AAA	TCT	GTC	TTT	GAT	GGC	GGA	GAC	GGT	624
166	Val	Asp	Gly	Val	Ser	Asp	His	Lys	Ser	Val	Phe	Asp	Gly	Gly	Asp	Gly	
167					200					205					210		
168	ACA	GTT	TAT	CAA	AAT	ATT	CAG	CAA	TTT	ATC	GAT	GAA	GGC	AAG	TGG	ATT	672
169	Thr	Val	Tyr	Gln	Asn	Ile	Gln	Gln	Phe	Ile	Asp	Glu	Gly	Lys	Trp	Ile	
170				215					220					225			
171	TCA	GGT	GAT	AAC	CAT	ACT	TTA	AGA	GAC	CCT	CAC	TAT	GTT	GAA	GAT	AAG	720
172	Ser	Gly	Asp	Asn	His	Thr	Leu	Arg	Asp	Pro	His	Tyr	Val	Glu	Asp	Lys	
173				230				235					240				
174	GGC	CAT	AAA	TAT	CTT	GTC	TTT	GAA	GCG	AAT	ACT	GGA	ACA	ACA	GAT	GGT	768
175	Gly	His	Lys	Tyr	Leu	Val	Phe	Glu	Ala	Asn	Thr	Gly	Thr	Thr	Asp	Gly	
176				245				250					255				
177	TAT	CAA	GGC	GAT	CAG	TCT	TTC	AAT	AAT	AAA	GCT	TAC	TAT	GGC	GGA	AGT	816
178	Tyr	Gln	Gly	Asp	Gln	Ser	Phe	Asn	Asn	Lys	Ala	Tyr	Tyr	Gly	Gly	Ser	
179				260			265			270					275		
180	GAC	GTC	TTC	TTC	CAG	AAT	GAA	AAA	AAT	AAA	CTG	CTT	CAA	AGT	CCT	AAA	864
181	Asp	Val	Phe	Phe	Gln	Asn	Glu	Lys	Asn	Lys	Leu	Leu	Gln	Ser	Pro	Lys	
182					280					285				290			
183	AAA	CAA	ATT	GCT	TCT	TTA	GCG	AAT	GGT	GCA	TTA	GGC	ATT	GTT	GAA	TTG	912
184	Lys	Gln	Ile	Ala	Ser	Leu	Ala	Asn	Gly	Ala	Leu	Gly	Ile	Val	Glu	Leu	
185				295					300					305			
186	GCC	GAT	GAC	TAT	ACA	GTG	AAA	AGT	GTT	ATG	AAA	CCA	TTA	GTC	GCA	TCA	960
187	Ala	Asp	Asp	Tyr	Thr	Val	Lys	Ser	Val	Met	Lys	Pro	Leu	Val	Ala	Ser	
188				310				315					320				
189	AAC	ACA	GTA	GCA	GAT	GAA	GTC	GAA	CGC	GCC	AAT	ATA	TTT	AAA	ATG	AAT	1008
190	Asn	Thr	Val	Ala	Asp	Glu	Val	Glu	Arg	Ala	Asn	Ile	Phe	Lys	Met	Asn	
191				325				330					335				
192	AAT	AAA	TGG	TAT	CTA	TTC	ACG	GAT	TCA	AGA	GGA	TCC	AAA	ATG	ACG	AGT	1056
193	Asn	Lys	Trp	Tyr	Leu	Phe	Thr	Asp	Ser	Arg	Gly	Ser	Lys	Met	Thr	Ser	
194				340			345				350				355		
195	GAT	GGA	ATT	AAC	GAC	AAA	GAT	GTT	TAT	ATG	CTA	GGG	CCC	GGA	GGC	GAC	1104
196	Asp	Gly	Ile	Asn	Asp	Lys	Asp	Val	Tyr	Met	Leu	Gly	Pro	Gly	Gly	Asp	
197					360					365				370			
198	TCC	TTA	AAT	GGC	CCA	CAC	AAC	CCG	ATA	AAT	GAA	ACT	GGA	CTT	GTA	TTG	1152
199	Ser	Leu	Asn	Gly	Pro	His	Asn	Pro	Ile	Asn	Glu	Thr	Gly	Leu	Val	Leu	
200				375					380					385			
201	AAC	ATG	AAT	CTT	GAC	CCT	GCT	GAT	CTC	ACA	CAC	ACT	TAC	TCT	CAT	TGC	1200
202	Asn	Met	Asn	Leu	Asp	Pro	Ala	Asp	Leu	Thr	His	Thr	Tyr	Ser	His	Cys	
203				390				395					400				
204	GGT	ATC	CCG	CAC	CCT	GAA	GGT	AAT	AAT	GTG	GTA	CTC	ACA	AGT	TAT	ATG	1248
205	Gly	Ile	Pro	His	Pro	Glu	Gly	Asn	Asn	Val	Val	Leu	Thr	Ser	Tyr	Met	

# RAW SEQUENCE LISTING PATENT APPLICATION US/09/986,682A

DATE: 02/21/2002  
TIME: 22:54:16

INPUT SET: S36770.raw

```

206          405          410          415
207 ACG AAT AGA GGC TTC TAT CCA GAA CAT CAC TCT CAC CTG CGG GAC AAG 1296
208 Thr Asn Arg Gly Phe Tyr Pro Glu His His Ser His Leu Arg Asp Lys
209 420          425          430          435
210 CTT GGG GTT AAT ATT AAA GGG TCT GAC ACA TCT GGA GGA GAA AAT AGT 1344
211 Leu Gly Val Asn Ile Lys Gly Ser Asp Thr Ser Gly Gly Glu Asn Ser
212          440          445          450
213 TCC GGA CAA GGA CAA TTC CCA 1365
214 Ser Gly Gln Gly Gln Phe Pro
215          455
216
217 (2) INFORMATION FOR SEQ ID NO: 5:
218 (i) SEQUENCE CHARACTERISTICS:
219 (A) LENGTH: 2408 base pairs
220 (B) TYPE: nucleic acid
221 (C) STRANDEDNESS: double
222 (D) TOPOLOGY: linear
223 (ii) MOLECULE TYPE: genomic DNA
224 (vi) ORIGINAL SOURCE:
225 (A) ORGANISM: Bacillus sp.
226 (C) INDIVIDUAL ISOLATE: V230 (FERM BP-5054)
227 (ix) FEATURE:
228 (A) NAME/KEY: signal peptide
229 (B) LOCATION: 1..360
230 (C) IDENTIFICATION METHOD: E
231 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 5:
232 CGGGGAAAAT ACTAGATTCC AATTGGCCAG ACTTCCCAGT TGGTGTAAGA GAAGAGTTTCG 60
233 GACTGCCAAT GCAGCTGTGC GTAAGAAAAC AGCTTACTCA TGAGCAATTA CTAGAAGAAT 120
234 TTCAAAGTC CTGGGATAAG GCCAAGTCCA CTTTGAAATA AACTTTTCAG CCTCTGTGTG 180
235 GGGGCTTTTT TGTTTTTTATT TATTTCAACT GCAAGTGGTC CATCCCCTAT ATCAATTTAA 240
236 GACGAAATTC TAATCAATCC ATGCCATCCC CAATAAACTC GTCCTCCTCT ATACTTTTAA 300
237 TTAATAAGAA ACTATCAAGA GCTTTCTTAT CAAATTCATA CATATCCAAG GAGGGAGACG 360
238 ATG AAC TTC AAA AGA TTG GCG AAA AAA GCA GCT GCC GTA ACC TTC AGG 408
239 Met Asn Phe Lys Arg Leu Ala Lys Lys Ala Ala Val Thr Phe Arg
240 -30 -25 -20
241 ACT GCT ATA TTA GTA GGA GCG GAC GGA CCG CAT ATT TTT GCG CAG CAA 456
242 Thr Ala Ile Leu Val Gly Ala Asp Gly Pro His Ile Phe Ala Gln Gln
243 -15 -10 -5
244 ATG AAC AGC GGG GAC TAC AAG GAA GAC TAT GGT TTT GCC CAT ATT ACA 504
245 Met Asn Ser Gly Asp Tyr Lys Glu Asp Tyr Gly Phe Ala His Ile Thr
246 1 5 10 15
247 CGC GCT GAC ATG CTA AAA ATT CCA GGA CAA CAA AAC AGT CCT CAA TTT 552
248 Arg Ala Asp Met Leu Lys Ile Pro Gly Gln Gln Asn Ser Pro Gln Phe
249 20 25 30
250 AAA GTG CCT CAA TTC AAT GCA TCA GCA ATC AAA AAC ATT GAT TCG GCA 600
251 Lys Val Pro Gln Phe Asn Ala Ser Ala Ile Lys Asn Ile Asp Ser Ala
252 35 40 45
253 AAA GGG TAT GAT AAG TCA GGC AAC TTA ATA GAT TTA GAT GTA TGG GAT 648
254 Lys Gly Tyr Asp Lys Ser Gly Asn Leu Ile Asp Leu Asp Val Trp Asp
255 50 55 60
256 AGC TGG CCA CTG CAA AAC GCT GAT GGT ACT GCG GCA AAT TAT CAT GGA 696
257 Ser Trp Pro Leu Gln Asn Ala Asp Gly Thr Ala Ala Asn Tyr His Gly
258 65 70 75 80

```

PAGE: 1

**SEQUENCE VERIFICATION REPORT**  
**PATENT APPLICATION US/09/986,682A**

DATE: 02/21/2002  
TIME: 22:54:16

***INPUT SET: S36770.raw***

Line

Error

Original Text

PAGE: 1

**SEQUENCE MISSING ITEM REPORT**  
**PATENT APPLICATION US/09/986,682A**

DATE: 02/21/2002  
TIME: 22:54:16

*INPUT SET: S36770.raw*

<< THERE ARE NO ITEMS MISSING >>

PAGE: 1

**SEQUENCE CORRECTION REPORT**  
**PATENT APPLICATION US/09/986,682A**

DATE: 02/21/2002  
TIME: 22:54:16

*INPUT SET: S36770.raw*

Line	Original Text	Corrected Text
226	(C) INDIVIDUAL ISOLATE: V230 (FERM BP-5054)	(C) INDIVIDUAL ISOLATE: V230 (FERM BP-5054)